[DRAFT] DETC2015-46982

DEVELOPMENT OF A CONSERVATIVE MODEL VALIDATION APPROACH FOR RELIABLE ANALYSIS

Min-Yeong Moon, K.K. Choi[*], Hyunkyoo Cho and Nicholas Gaul
Department of Mechanical & Industrial Engineering College of Engineering, The University of Iowa, Iowa City, IA 52242, USA
minyeong-moon@uiowa.edu, kyung-choi@uiowa.edu, hyunkyoo-cho@uiowa.edu and nicholas-gaul@ramdosolution.com
*Corresponding author

David Lamb, David Gorsich
US Army RDECOM/TARDEC, Warren, MI 48397-5000, USA
david.lamb@us.army.mil, david.j.gorsich.civ@mail.mil

ABSTRACT

Simulation models are approximations of real-world physical systems. Therefore, simulation model validation is necessary for the simulation-based design process to provide reliable products. However, due to the cost of product testing, experimental data in the context of model validation is limited for a given design. When the experimental data is limited, a true output PDF cannot be correctly obtained. Therefore, reliable target output PDF needs to be used to update the simulation model. In this paper, a new model validation approach is proposed to obtain a conservative estimation of the target output PDF for validation of the simulation model in reliability analysis. The proposed method considers the uncertainty induced by insufficient experimental data in estimation of predicted output PDFs by using Bayesian analysis. Then, a target output PDF and a probability of failure are selected from these predicted output PDFs at a user-specified conservativeness level for validation. For validation, the calibration parameter and model bias are optimized to minimize a validation measure of the simulation output PDF and the conservative target output PDF subject to the conservative probability of failure. For the optimization, accurate sensitivity of the validation measure is obtained using the complex variable method (CVM) for sensitivity analysis. As the target output PDF satisfies the user-specified conservativeness level, the validated simulation model provides a conservative representation of the experimental data. A simply supported beam is used to carry out the convergence study and demonstrate that the proposed method establishes a conservatively reliable simulation model.

KEYWORDS

Conservative Model Validation, Insufficient Experimental Data, Conservative Output PDF, Bayesian Analysis, Model Bias, Calibration Parameter

1. INTRODUCTION

Computer simulation plays increasingly important roles in various engineering design projects with rapid increase of computational power. Accordingly, simulation-based design enables engineers to design a product with less time and cost by reducing the number of hardware prototyping and testing of the designed product. However, an accurate, safe, and reliable simulation model is not easy to obtain due to the approximate imitation of real systems including idealization and simplification. As a result, model validation has become an important topic. Model validation can be defined by two parts: (1) updating a simulation model by utilizing mathematical formulations with given experimental data and (2) the assessment of the accuracy of the simulation model by comparing simulation result and experimental data [1-4]. In this paper, we focus on updating the simulation model to build a reliable simulation-based design given few experimental data.

Various model updating techniques have been developed over the last decade by quantifying the source of uncertainty involved in computer simulation such as calibration parameters and model bias (or model discrepancy). Methods to consider the source of uncertainty can be categorized into two main strategies: the Bayesian approach [5-8] and maximum likelihood estimation (MLE) [9-11].

Kennedy and O'Hagan proposed a Bayesian approach to obtain posterior distributions of calibration parameter and model

Report Documentation Page

1. REPORT DATE **29 APR 2015**	2. REPORT TYPE	3. DATES COVERED **00-00-2015 to 00-00-2015**

4. TITLE AND SUBTITLE **DEVELOPMENT OF A CONSERVATIVE MODEL VALIDATION APPROACH FOR RELIABLE ANALYSIS**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army RDECOM-TARDEC,6501 E. 11 Mile Road,Warren,MI,48397-5000**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
See Report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

bias function [5]; and their research has been applied and extended by many researchers. Hidgdon followed the Bayesian model validation approach and extended it to solve a multivariate output problem [6]. Arendt et al. applied the Bayesian approach for both single and multiple responses and discussed the identifiability of calibration parameter and model bias. They found that multiple responses can decrease the standard deviations of posterior distributions of calibration parameter and model bias compared to the results from single response [7, 8]. One limitation of the aforementioned Bayesian model validation is that the unknown calibration parameter used in the simulation model is treated as a constant. Even though the posterior distribution of calibration parameter can be obtained, a user has to choose one value to update simulation model. However, the calibration parameters to be adjusted in an engineering problem such as material properties can vary randomly due to manufacturing variability; thus, the calibration parameter must not be fixed in the simulation model. Another shortcoming is that existing Bayesian model validations do not guarantee that the predicted output PDF leads to safe design because predicted output PDF is based on few test data and prior knowledge without considering a conservative concept. Therefore, the predicted probability of failure obtained from the existing Bayesian model validations could be less than true probability of failure.

At the same time, there have been research efforts to maximize the agreement between the simulation model and experimental data using MLE. Loeppky et al. suggested MLE to find the fixed value of calibration parameter and bias function assumed to follow Gaussian process for deterministic problem [9]. Xiong et al. realized that a fixed calibration parameter is not consistent with physical experiments, so they applied MLE to estimate the statistical parameters of randomly varying unknown parameters [10]. This approach to calibrate statistical parameters has been followed by Youn et al. [11]. However, typically, there is only a small amount of experimental data available in the context of model validation due to the expensive cost of full-scale product testing. Consequently, the lack of experimental data leads to inaccurate likelihood calculation. Thus, output probability density function (PDF) estimated via MLE may not be accurate. Moreover, as the MLE method cannot provide a conservative measure, the updated simulation based on MLE could underestimate a probability of failure so that the simulation-based design would be unreliable. In this situation, rather than trying to match the simulation model to the small amount of experimental data, it is more desirable to obtain a conservative simulation model for reliable design even with limited experimental data.

Very little research has taken into account the conservative or safe design for model validation of computer simulation. However, a conservative or safe design has been achieved in many fields by using various conservative measures such as a safety factor, a conservative material properties [12]. As for the reliability problem, a conservative design can be achieved by estimating a larger probability of failure than the true value. Picheny et al. recognized the danger of underestimating the

probability of failure due to the limited number of data, so they developed a conservative estimation of probability of failure using a bootstrap method assuming the data distribution type is known [13]. In the same context, several reliability-based design optimization (RBDO) approaches have been developed to obtain a conservative design in compensation for the lack of input data. Youn and Wang sought the worst case from distribution of reliability, which is beta distribution, and defined it as target reliability for RBDO [14]. Noh et al. proposed the confidence level of input statistical model with adjusted standard deviation and correlation coefficient, which fully covers the target reliability region for RBDO [15, 16]. Cho used a confidence level of distribution of probability of failure as a probabilistic constraint for confidence-based RBDO [17].

In this paper, we emphasize that the validated simulation model should lead an optimum conservative design so as to acquire a safe design even with a small number of full-scale product testing. To ensure that, a conservative estimations of target output PDF and probability of failure are achieved by quantifying the uncertainty caused by the limited test using Bayesian analysis in this paper. Once the conservative target output PDF at a user-specified conservativeness level is obtained, model bias and calibration parameters are characterized through a conservative model-updating optimization process according to the given design. The characterized model bias and calibration parameters update the simulation model. Then, the updated simulation model will be appropriately conservative and trustworthy for RBDO, which prevents the risk of underestimation of probability of failure.

The remainder of the paper is outlined as follows. Section 2 describes the various sources of uncertainty involved in computer simulation. In Section 3, the proposed conservative model validation is briefly compared to the conventional model validation approach. Section 4 describes how to account for the uncertainty of lack of experimental data and estimate conservative target output PDF and probability of failure using Bayesian analysis. Section 5 introduces the conservative model-updating optimization by characterizing calibration parameters and model bias. The proposed method is applied to a simply supported beam and the results demonstrate that the updated simulation model provides conservative fit to true one in Section 6.

2. VARIOUS SOURCES OF UNCERTAINTIES

In Ref. 5, various sources of uncertainties in the use of computer simulation have been identified. The first source is parametric variability which is the randomness of the input variable. The thickness of a steel plate that can vary randomly within a tolerance could be an example. It might not be exactly designed and constructed due to the manufacturing error. Another source is parameter uncertainty, which comes from input variables whose exact values (or probabilistic distribution) are unknown. For example, material properties may vary in the physical experiments due to their variability. In this paper, statistical properties of unknown parameter are defined as calibration parameters. In addition, structural uncertainty,

which is model bias or discrepancy, refers to the fundamental inability to reproduce the real-world because of simplification and idealization. Furthermore, there is numerical uncertainty which is a numerical error in the implementation of the computer simulation as well. However, simulation is often assumed to be numerical error-free. Another source is experimental uncertainty (or observation error) which is an error in measuring the experimental responses. This can be possibly noticed by repeating measurements many times. Among those various uncertainties, parameter uncertainty, input variability and model bias are concerned in this paper. In addition, the uncertainty also comes from insufficient experimental data. Hence, the uncertainty induced by the lack of experimental data is taken into consideration in the proposed conservative model validation approach.

A model-updating formulation that combines a simulation model and incorporates experimental data as well as various uncertainties is generally outlined as

$$\mathbf{y}^{e}\left(\mathbf{x}\right) = \mathbf{y}^{m}\left(\mathbf{x}; \theta\right) + \delta\left(\mathbf{x}\right) + \varepsilon\left(\mathbf{x}\right) \qquad (1)$$

where $\mathbf{x}$ is an input random variable vector, θ is an uncontrollable input random variable to be calibrated (e.g., material property and friction coefficient), $\varepsilon(\mathbf{x})$ is experimental error which is not considered in this paper, $\delta(\mathbf{x})$ represents model bias, $\mathbf{y}^{m}(\mathbf{x}; \theta)$ is the simulation model, and $\mathbf{y}^{e}(\mathbf{x})$ is the experimental data. The simulation model is less expensive than the actual experiments. Hence, it is assumed that the simulation model can be evaluated as many as needed. In contrast, only a limited number of experimental data, $\mathbf{y}^{e}_{1}, \mathbf{y}^{e}_{2}, \ldots, \mathbf{y}^{e}_{N}$, can be collected at given input setting $\mathbf{x}$. Due to the uncertainty induced by lack of experimental data, a true output PDF cannot be exactly determined; and a predicted output PDF is uncertain. What makes it more challenging is that the input setting $\mathbf{x}$ corresponding to the experimental data may not be known. In this paper, we address the case when the input setting $\mathbf{x}$ corresponding experimental data is not known. In following sections, all of the possible sources of uncertainty are identified and characterized.

3. PROPOSED CONSERVATIVE MODEL VALIDATION FRAMEWORK

Conventional model validation approaches focus on updating a simulation model directly based on physical experimental data of which only a few are available. Once the simulation model is validated, it is then applied for a design purpose. Model updating can be viewed as the mathematical procedure that characterizes calibration parameter and model bias. However, when only limited resources for experimental data are available, conventional model validation does not assure that updated simulation model is reliable. Consequently, the actual test of a product that is designed based on the validated simulation model could fail.

In this paper, a conservative model validation approach is proposed with an emphasis on assuring reliable model-updating

for design optimization. The flowchart of the proposed approach is given in Fig. 1. There are two main differences between the developed method and the conventional approaches. First, the proposed model validation investigates and quantifies the uncertainty induced by the lack of physical experimental data for the purpose of conservativeness. Next, the proposed method estimates the distribution of probability of failure. At a user-specified conservativeness level of probability of failure, an output PDF is obtained for an acceptable conservativeness, which is called *conservative target output PDF* in this paper. After that, model-updating optimization is performed to match the simulation output PDF to the conservative target output PDF by satisfying the user-specified conservativeness level. Then, the updated simulation model provides a conservative set of calibration parameter and bias.
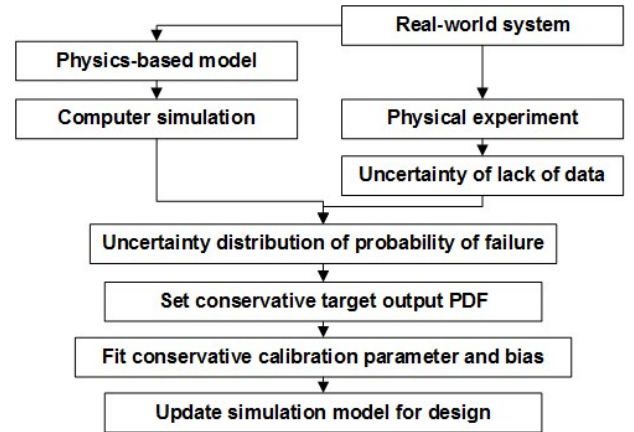

Figure 1. Flowchart of Proposed Model Validation Approach

4. QUANTIFICATION OF UNCERTAINTY INDUCED BY INSUFFFICIENT EXPERIMENTAL DATA

In this section, we consider how to model and quantify the uncertainty due to insufficient experimental data. With limited experimental data, a predicted output PDF becomes uncertain and subjective. As a result, a validated model based on the uncertain and subjective output PDF becomes uncertain as well. Accordingly, an uncertainty also exists in the probability of failure, which is calculated based on the validated model. Finally, the uncertainty due to the limited data propagates to the uncertainty of the probability of failure. Therefore, the probability of failure calculated using the validated model cannot indicate the reliability of the physical system correctly. Hence, in this paper, we estimate the distribution of probability of failure which involves the uncertainty due to the limited experimental data. The following sections explain the developed method in detail.

4.1 Distribution of probability of failure

As mentioned earlier, we address the fact that experimental data $\mathbf{y}^{e}$ is small. The uncertainty caused by insufficient data leads to the uncertainty of a predicted output PDF and consequently uncertainty of a probability of failure. To quantify

the uncertainty of the predicted output PDF, parameters that can represent various statistical information about the predicted output PDF need to be selected. In this paper, the first four statistical moments of the predicted output PDF have been chosen as the parameters we are interested in. Uncertainty of the predicted output PDF is characterized using the uncertainties of these statistical moments.

Bayesian analysis, which enables quantifying the uncertainty distribution of the parameters (the first four statistical moments in this paper), has been used to model the uncertainty of a probability of failure. A joint PDF of the probability of failure and the four statistical moments can be obtained using the Bayesian formulation as follows:

$$f\left(p_F, \boldsymbol{\mu} \mid \mathbf{y}^e\right) = f\left(p_F \mid \boldsymbol{\mu}, \mathbf{y}^e\right) P\left(\boldsymbol{\mu} \mid \mathbf{y}^e\right) \qquad (2)$$

where $\boldsymbol{\mu} = [\mu, \sigma^2, \beta_1, \beta_2]$ represents four statistical moment vector, $\mathbf{y}^e$ is available experimental data vector and p_F is the estimated probability of failure. The posterior distribution of four statistical moments, $P(\boldsymbol{\mu}|\mathbf{y}^e)$, can be calculated as the product of likelihood function and prior function as:

$$P\left(\boldsymbol{\mu} \mid \mathbf{y}^e\right) = L\left(\mathbf{y}^e; \boldsymbol{\mu}\right) P\left(\boldsymbol{\mu}\right). \qquad (3)$$

The prior function $P(\boldsymbol{\mu})$ in Eq. (3) is the product of priors of four statistical moments assuming that they are independent of each other. Information about four statistical moments of the computer simulation output PDF has been used to represent the prior function. Likelihood function $L(\mathbf{y}^e; \boldsymbol{\mu})$ is the product of the predicted output PDF values at each experimental data point given four statistical moment vector $\boldsymbol{\mu}$. The predicted output PDF given four statistical moments can be uniquely estimated using various methods. In this paper, the Pearson system [18] has been used to evaluate the likelihood function. The Pearson system is a non-parametric method that can approximate PDF given four statistical moments. It categorizes the PDF into seven different types of continuous distribution depending on the four moments, which can cover various shapes of distribution. Accordingly, the probability of failure can be uniquely determined given four statistical moments. Therefore, in Eq. (2), the conditional PDF of probability of failure given four statistical moments and experimental data becomes a Dirac measure as

$$f\left(p_F \mid \boldsymbol{\mu}, \mathbf{y}^e\right) = \delta\left[p_F - p_F\left(\boldsymbol{\mu} \mid \mathbf{y}^e\right)\right]. \qquad (4)$$

Consequently, the probability, which is the integration of Eq. (4), that p_F equals $p_F(\boldsymbol{\mu}|\mathbf{y}^e)$ becomes one; and the probability of getting any other values is zero. The marginal PDF of the probability of failure can be calculated using Monte Carlo Simulation (MCS) integration as

$$f_{p_F}\left(p_F \mid \mathbf{y}^e\right) = \int_{\Omega_{\boldsymbol{\mu}}} f\left(p_F \mid \boldsymbol{\mu}, \mathbf{y}^e\right) P\left(\boldsymbol{\mu} \mid \mathbf{y}^e\right) d\boldsymbol{\mu}$$

$$= \frac{1}{nMCS} \sum_{i=1}^{nMCS} \left[f\left(p_F \mid \boldsymbol{\mu}^{(i)}, \mathbf{y}^e\right)\right]$$

$$= \frac{1}{nMCS} \sum_{i=1}^{nMCS} I_{p_F\left(\boldsymbol{\mu}^{(i)}|\mathbf{y}^e\right)}\left(p_F\right) \qquad (5)$$

$$where \ I_{p_F\left(\boldsymbol{\mu}^{(i)}|\mathbf{y}^e\right)}\left(p_F\right) = \begin{cases} 1 & if \ p_F = p_F\left(\boldsymbol{\mu}^{(i)} \mid \mathbf{y}^e\right) \\ 0 & if \ p_F \neq p_F\left(\boldsymbol{\mu}^{(i)} \mid \mathbf{y}^e\right) \end{cases}$$

where $nMCS$ is the number of MCS samples and $\boldsymbol{\mu}^{(i)}$ is the i-th realization of the four statistical moment vector. Because the posterior distribution in Eq. (3) cannot be derived analytically, the realizations of four statistical moments are generated using the Markov Chain Monte Carlo sampler in accordance with the probability in Eq. (3).

4.2 Conservative target output PDF

To obtain a reliable design, the estimated probability of failure using the simulation model should be equal to or larger than the true probability of failure. However, if the probability of failure is too much overestimated, the optimum cost of the simulation-based RBDO will unnecessarily increase. Therefore, the simulation model should be fit to be appropriately conservative from the designer's viewpoint. To tackle this issue, a conservativeness level is introduced in this paper, which can control how conservative the design will be. Once the distribution of probability of failure is obtained as explained in the previous section, the quantile of the distribution can be interpreted as the conservativeness level as shown in Fig. 2. For example, 90% quantile value of distribution of probability of failure indicates 90% conservativeness level. Thus, the probability of failure at 95% conservativeness level is higher than the one at 90% conservativeness level. The probability of failure at the user-specified conservativeness level is selected as *the conservative probability of failure*, and the corresponding predicted output PDF will be chosen as a *conservative target output PDF*. Even though the conservative target output PDF may not represent the data or the true output PDF accurately, it will provide an appropriately conservative fit to the experimental data. Moreover, the conservative target output PDF will converge to the true output PDF as the amount of experimental data increases. This convergence study will be demonstrated in Section 6.
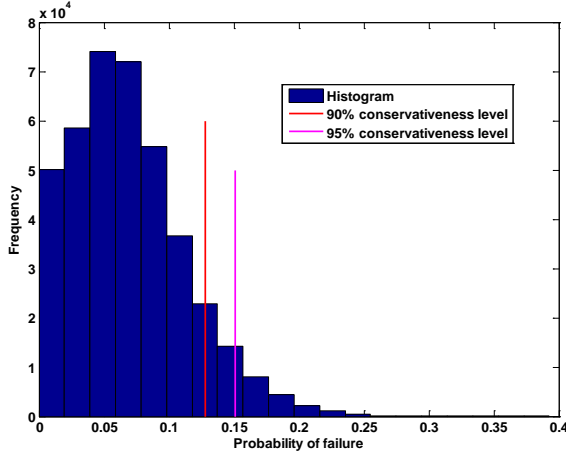
Figure 2. Conservativeness Level: Quantile of Probability of Failure

5. CONSERVATIVE MODEL-UPDATING OPTMIZATION

5.1 Formulation of conservative model-updating optimization

Model-updating optimization minimizes a validation measure which is defined as the distance between two distributions: the simulation output PDF and conservative target output PDF. In order to assure a user-specified conservativeness level, the conservative probability of failure obtained from Section 4 must be maintained. The mathematical formulation of conservative model-updating optimization is shown as:

$$\text{minimize } H(\mathbf{d}) = \frac{1}{2}\int_{-\infty}^{\infty}\left(\sqrt{p\left(g\left(\mathbf{x}\right);\mathbf{d}\right)} - \sqrt{q\left(g\right)}\right)^2 dg$$

$$= 1 - \int_{-\infty}^{\infty}\sqrt{p\left(g\left(\mathbf{x}\right);\mathbf{d}\right)q\left(g\right)}dg \qquad (6)$$

$$\text{subject to } p_F\left(\mathbf{x};\mathbf{d}\right) = \text{P}\left(p\left(g\left(\mathbf{x}\right) > g_{\lim};\mathbf{d}\right)\right) = p_F^{con}$$

$$\text{where } \mathbf{d}^L \le \mathbf{d} \le \mathbf{d}^U, \quad \mathbf{d} \in \mathbb{R}^4 \text{ and } \mathbf{x} \in \mathbb{R}^{NR}$$

and $p(g(\mathbf{x});\mathbf{d})$ is the simulation output PDF, $q(g)$ is the conservative target output PDF obtained from Section 4, $\mathbf{x}$ is an input random variable vector, $\mathbf{d}$ is a optimization variable vector that includes statistical calibration parameters and statistical properties of bias function, $g(\mathbf{x}) > g_{\lim}$ defines a failure region, NR is the number of input random variables and p_F^{con} is the conservative probability of failure obtained from Section 4. The Hellinger similarity $H(\mathbf{d})$, which is a validation measure, is used as an objective function to be minimized. Here, it is noted that the model-updating optimization process in Eq. (6) does not try to exactly match the conservative target output PDF. Accordingly, the optimized statistical properties of bias and calibration parameters may not be true. The more important thing is that the simulation model must have the same probability of failure as the conservative probability of failure; thus, an equality constraint is used in Eq. (6). That is, we do not use an inequality constraint lest it will yield the validated model that provides too conservative design if we use it. In the meantime, the simulation model will have the closest output distribution to

the conservative target output PDF by minimizing the Hellinger similarity in the optimization process. The optimum results characterize statistical calibration parameters and bias function by satisfying a user-specified conservativeness level. As the equality constraint in Eq. (6) satisfies the conservative target probability of failure, the obtained statistical calibration parameter and the model bias provide conservative fit to the experimental data and prevents the underestimation of the probability of failure.

5.2 Sensitivity using Complex Variable Method (CVM)

The sensitivity of the validation measure (i.e., cost function) in Eq. (6) requires the sensitivity of the simulation output PDF $p(g(\mathbf{x});\mathbf{d})$. However, the simulation output PDF cannot be analytically evaluated and thus MCS is used. As a result, a certain amount of numerical error (MCS error) occurs in evaluating the validation measure due to the MCS error. This numerical error severely hinders the accurate evaluation of the sensitivity of the validation measure using the finite difference method (FDM). As the FDM sensitivity of a validation measure with respect to j-th optimization variable $\mathbf{d}_j$ can be formulated as Eq. (7), it has both truncation and round-off error inherently.

$$\frac{\partial H}{\partial \mathbf{d}_{j\ FDM}} \approx \frac{H\left(\mathbf{d}_j + \Delta\mathbf{d}_j\right) - H\left(\mathbf{d}_j - \Delta\mathbf{d}_j\right)}{2\Delta\mathbf{d}_j}. \qquad (7)$$

Hence, an appropriate finite difference step size $\Delta\mathbf{d}_j$ is very difficult to obtain, so the FDM will suffer the numerical error. In this paper, the complex variable method (CVM) [19] has been used to improve the accuracy of the sensitivity. The CVM sensitivity of the validation measure with respect to j-th optimization variable $\mathbf{d}_j$ can be defined as the imaginary part of a complex-valued function in Eq. (8).

$$\frac{\partial H}{\partial \mathbf{d}_{j\ CVM}} \approx \frac{\text{Im}\left[H\left(\mathbf{d}_j + i\Delta\mathbf{d}_j\right)\right]}{\Delta\mathbf{d}_j}. \qquad (8)$$

Unlike FDM, CVM has only truncation error, which can be alleviated by selecting sufficiently small step size. Therefore, a stable and accurate sensitivity can be achieved. One limitation of CVM is that performance measure analysis should be able to handle a complex variable. However, this limitation is resolved by using a surrogate model for the performance measure. The accuracy of the sensitivity using CVM is demonstrated in Section 6.

6. NUMERICAL EXAMPLES

In this section, the accuracy of the sensitivity of the validation measure using the CVM in Eq. (6) is verified. In addition, the conservative model validation is demonstrated using the simply supported beam example and is compared to the model validation using MLE. It is noted that the conservative model validation prevents the underestimation of the probability

of failure whereas MLE cannot. As the amount of experimental data increases, it is shown that the updated simulation model by the conservative model validation converges to the experimental data.

6.1 Sensitivity verification

The example used to verify the sensitivity of the validation measure is a 3D solid cantilever beam model (one 8-noded element) as shown in Fig. 3 [20]. The thickness in the y-direction at the free end (right end in Fig. 3) of the beam model is a random design variable, X in Fig. 3, which follows $N\sim(1, 0.05^2)$. The mean μ_θ and standard deviation σ_θ of Poisson ratio θ are statistical calibration parameters. Hence the optimization variable vector $\mathbf{d}$ of model-updating optimization in Eq. (6) is $[\mu_\theta, \sigma_\theta]$. The Young's modulus is 200,000 lb/in^2 and the applied pressure load is 20 lb/in^2. The output g in Eq. (6) is the von-Mises stress of the beam at element centroid and failure occurs when the von-Mises stress is larger than 50ksi. The validation measure $H(\mathbf{d})$ in Eq. (6) is defined as the distance between two PDFs of von-Mises stress. The first PDF $p(g(X);\mathbf{d})$ is the simulation output PDF which is obtained using finite element (FE) analysis with the random design variable X and Poisson ratio θ which follows $N(\mu_\theta, \sigma_\theta^2)$. The other one $q(g)$ is the target von-Mises stress PDF obtained at 95% quantile from the distribution of the probability of failure given ten points of experimental data. The data is generated by adding the hypothetical bias function exp(1.1X) to the FE analysis with the true distribution of the Poisson ratio $N\sim(2.3, 0.08^2)$.

The sensitivity of the validation measure with respect to the statistical calibration parameter μ_θ is computed for different perturbation sizes as shown in Table 1. Sensitivities of the validation measure are calculated at $[\mu_\theta, \sigma_\theta] = [2, 0.1]$. It is found that central FDM (CFDM) is extremely sensitive to the perturbation size while CVM is stable for any perturbation size. Thus, it can be seen that appropriate perturbation size of CFDM is hard to find. When the perturbation size is 7%~10%, the sensitivity using CFDM with respect to μ_θ is close to the result from CVM. Therefore, it implies that CVM provides accurate and stable sensitivity.
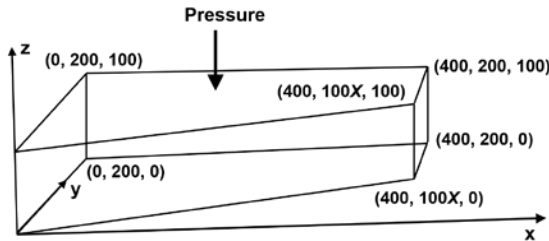

Figure 3. 3D Eight-node Cantilever Beam Model

Table 1. Comparison of Sensitivity of Validation Measure

Perturbation size	$dH / d\mu_\theta$		
	CVM (A)	CFDM (B)	Agreement (A/B×100)
0.001%	−0.06883	22.40762	-0.31%
0.01%	−0.06868	−0.81769	8.40%
0.1%	−0.06876	−0.41942	16.39%
1%	−0.06873	−0.08002	85.89%
4%	−0.06870	−0.06368	107.88%
7%	−0.06867	−0.06633	103.53%
10%	−0.06855	−0.06893	99.45%

6.2 Conservative model validation: simply supported beam

To demonstrate the proposed conservative model validation approach, a simply supported beam [7, 8] shown in Fig. 4 is considered as a numerical example in this section. To test the developed method in this paper, the original beam problem in Refs. 7 and 8 has been modified to have random design variables. The modified beam has a fixed length (2m), while its rectangular cross-section has random width and height. In addition, the static load, which is applied at the center of the beam, is a random parameter as well. Young's modulus is an unknown input variable whose statistical parameters are calibration parameters. The input random variables used in this example are defined in Table 2. The response of this simply supported beam example is the deflection at midpoint of the beam.

As shown in Fig. 5, two different stress-strain curves are used to stand for a simulation model and a physical experiment, respectively. Firstly, the left curve in Fig. 5 represents a simplified material (linear hardening for plastic region $\boldsymbol{\sigma}=\boldsymbol{\sigma}_Y+A\varepsilon$, where $\boldsymbol{\sigma}_Y$=225MPa and A=4100MPa). The surrogate response for the FE model using the simplified material has been created using the Dynamic Kriging (DKG) method [21, 22] and treated as a simulation model. It has been demonstrated that the DKG is the most accurate metamodeling method by having the smallest error compared to other metamodels [23, 24]. Secondly, the right curve in Fig. 5 indicates more realistic material (a power law [25] for plastic region with $\boldsymbol{\sigma}=\boldsymbol{\sigma}_Y+C\varepsilon^n$, where $\boldsymbol{\sigma}_Y$=225MPa, n=0.5 and C=2068MPa). The FE model using this more realistic material is treated as the true physical model that represents experiments.

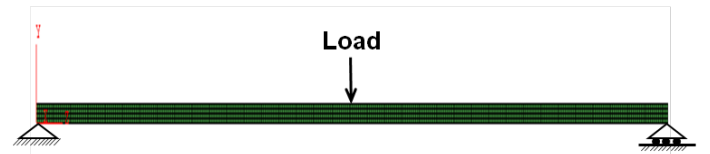Figure 4. Simply Supported Beam Model

Table 2. Information of Input Random Variables

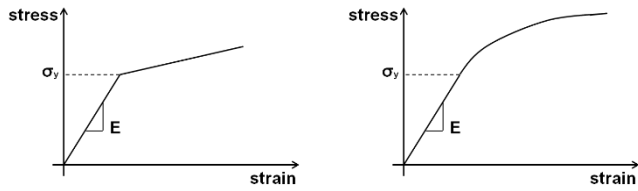| Input Random Variable	Distribution Type	Mean	Standard Deviation
Load (N)	Normal	1900	40
Width (mm)	Normal	52.5	1.6
Height (mm)	Normal	20	0.5
True Young's modulus (GPa)	Normal	206	4
Initial Guess of Young's modulus (GPa)	Normal	220	8

Figure 5. Stress-strain Curve (left: simulation model, right: experiments)

The simulation output PDF using the initial guess of Young's modulus and the true output PDF using the true distribution of Young's modulus are illustrated in Fig. 6. Failure occurs when the deflection at midpoint is larger than 51.5mm as shown in Fig. 6. In comparison with the true probability of failure (11.09%), it can be noted that the estimated probability of failure using the simulation model (3.24%) is underestimated at the current design. This implies that the simulation-based design could fail to the actual product testing due to the underestimated probability of failure. Thus, it calls for the conservative model validation method.
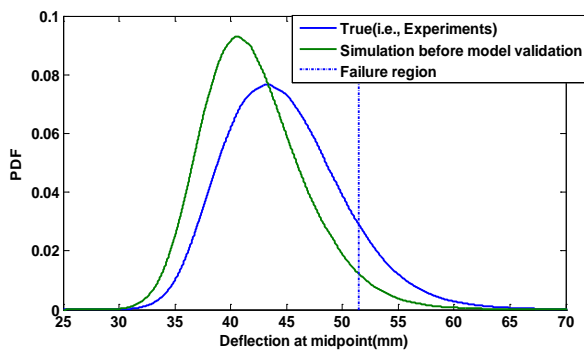

Figure 6. True and Simulation Output PDFs

In this example, we assume that the simulation output PDF is available because simulation responses using the Dynamic Kriging model are cheap and can be obtained as many as required. However, the true output PDF in Fig. 6 cannot be obtained in reality; typically only a small amount of data is available for the model validation due to the expense of full-scale

product testing. Therefore, the true probability of failure is usually unknown. To test if the developed model validation can tackle this issue, five experimental data points, which is a very small number, were randomly drawn from the true output PDF. Only the selected five data points are used for the model validation to demonstrate that the updated simulation model provides a conservative fit to the true output PDF.

Additionally, the result of the proposed model validation is compared to that from the model validation using MLE. For both methods, mean and standard deviation of Young's modulus are the statistical calibration parameter. The model bias $\delta(\mathbf{x})$ in Eq. (1) is assumed to follow normal distribution. The means and standard deviations of Young's modulus and model bias are four optimization variables in the model-updating optimization process of Eq. (6). Model validation using MLE does not include Bayesian analysis in Section 4 and uses only given data, without considering the uncertainty due to the lack of data. The conservative model validation using the proposed conservative model validation has been carried out using two different conservativeness levels of 90% and 95%. Model validations using both methods (the conservative model validation and MLE) have been performed 10 times for different datasets of five data points (which were drawn from the true output PDF).

The estimated probabilities of failure using the updated simulation models are shown in Table 3. As for the model validation with MLE, eight datasets out of ten (underlined in Table 3) underestimate the probabilities of failure. On the other hand, the conservative model validation (I) using 90% conservativeness level leads to the conservative design except for datasets 2, 7 and 8. When the conservativeness level is large (95%), the conservative model validation (II) leads to a safe design except for dataset 8, but maybe overly conservative overall compared to the true value of probability of failure. Even though the updated simulation model using the conservative model validation may not be exactly accurate, it prevents the danger of underestimating the probability of failure. Therefore, the conservative model validation is necessary when a small amount of experimental data is provided.

Output PDFs of the updated simulation model using both methods are shown in Fig. 7. Three datasets (1, 8 and 10) are used to present the updated simulation output PDFs. As shown in Fig. 7(a), dataset 1 is negatively skewed whereas the true output PDF is positively skewed. It is noted that this biased data leads to overly conservative results with limited data. Hence, the validated simulation model obtained from both methods overestimates the probability of failure. Dataset 8 given in Fig. 7(b) is extremely positively skewed; four out of five data points are clustered around the left tail of the true output PDF. In this situation, the probability of failure obtained from the model validation using MLE is merely 1.5%. Even the conservative model validation 2 using 95% conservativeness level result is not conservative enough (10.06%) compared to the true probability of failure (11.09%). However, it is clearly shown that the conservative model validation tends to be safer than the model validation with MLE. As shown in Fig. 7(c), dataset 10 is sparsely located and positively skewed. The
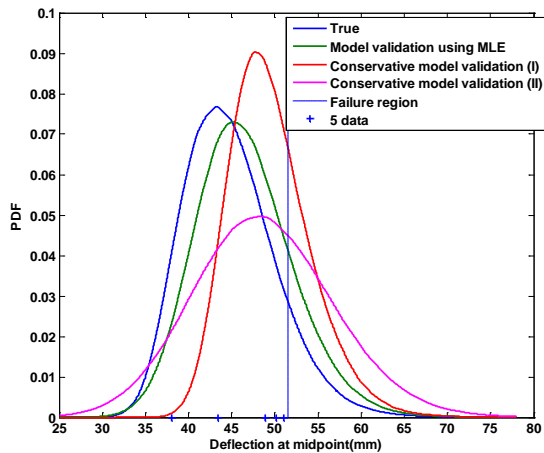
updated simulation model using MLE is less conservative (4.92%) than the true one, so the actual testing of a product designed based on the simulation model would not be safe. On the other hand, the conservative model validation provides a reliable simulation model by estimating the probability of failure conservatively (12.29% or 16.24%). Therefore, based on the model validation results under different data sets, it indicates that obtained experimental data must be appropriate and unbiased. If given small amount of experimental data is biased and does not represent the true output PDF well, even the conservative model validation falls into the underestimated probability of failure like data set 8.

Table 3. Estimated Probability of Failure of Updated Simulation Model under Different Data Sets
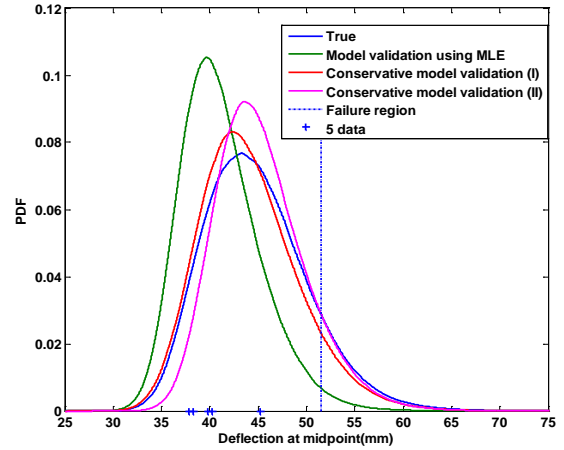
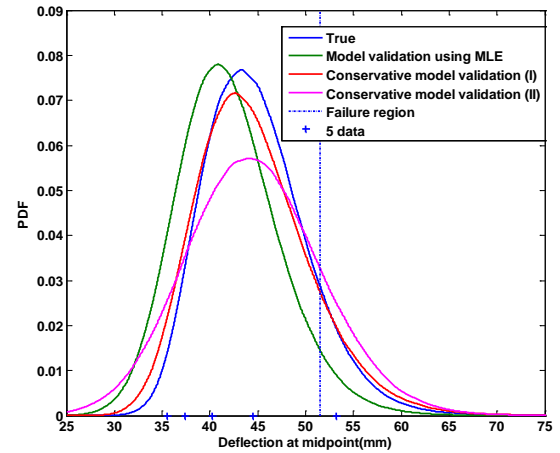| Dataset	Conservative model validation (I)[a]	Conservative model validation (II)[b]	Model validation using MLE	True P_F
1	29.87%	35.52%	18.18%	11.09%
2	9.82%	12.93%	4.10%	11.09%
3	11.51%	14.45%	4.70%	11.09%
4	18.90%	22.35%	9.18%	11.09%
5	21.20%	26.19%	8.72%	11.09%
6	33.29%	39.90%	18.72%	11.09%
7	8.59%	11.40%	4.03%	11.09%
8	7.40%	10.06%	1.50%	11.09%
9	28.83%	43.48%	8.84%	11.09%
10	12.29%	16.24%	4.92%	11.09%

[a] 90% conservativeness level
[b] 95% conservativeness level

(a) Dataset 1 – Negatively Skewed Data (skewness: -0.6854)

(b) Dataset 8 – Extremely Positively Skewed Data (skewness: 1.0651)

(c) Dataset 10 – Positively Skewed Data (skewness: 0.7583)

Figure 7. Comparison with Conservative Model Validation and MLE under Various Datasets

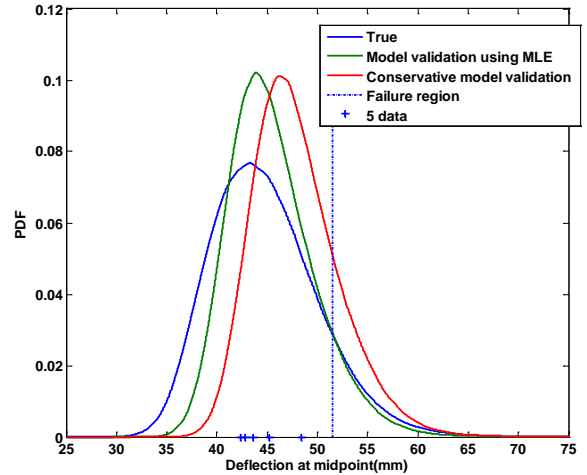6.3 Convergence study: the effect of experimental data size

In Section 6.2, we showed that the conservative model validation can provide an overall conservative fit of the true output PDF even with very limited data. On the other hand, the model validation with MLE has a high risk of underestimated probability of failure. However, certain biased data sets seems overly conservative when the conservative model validation as described in Section 6.2 is used. Hence, it is questionable whether it is indeed overly conservative or it is inevitable due the limited number of data that are biased. In this section, the amount of experimental data of the simply supported beam example is increased up to 4,000. Ten different amounts of data (5, 10, 50, 100, 300, 500, 1000, 2000, 3000 and 4,000) are considered to test the effect of experimental data size. Conservative model validation in this convergence test consistently used 90% conservativeness level. The estimated probability of failure of the updated simulation model with different data sizes is shown in Table 4. It is found that the

estimated probability of failure from both methods converges to the true probability of failure as the experimental data size increases. It can be noticed that the conservative model validation in the case of 4000 experimental data size is still not enough to be close to the true probability of failure. Theoretically, the posterior distribution of probability of failure is nothing but Dirac Delta measure when infinite experimental data size is available; however, only finite size of experimental data can be numerically used. Therefore, the conservative model validation using 90% conservativeness level tends to be conservative (11.95%) than the true value (11.09%) while the output PDF at posterior mean (11.42%) of distribution of probability of failure estimates closer to the true value.
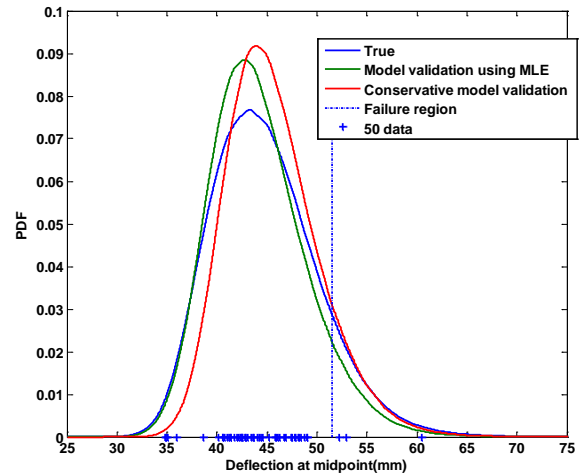
In addition, Fig. 8 shows that the updated simulation output PDF obtained from both methods converges to the true output PDF as the experimental data size increases. Therefore, it can be seen that both conservative model validation and MLE can provide an accurate simulation output PDF with a large amount of experimental data. Also, it can be found that the overly conservative results in Section 6.2 are due to the extremely small amount of data. Furthermore, in reality, actual product testing is extremely expensive in order to get a large amount of experimental data. Thus, accurate simulation output PDF cannot be obtained with a small amount of experimental data. In this paper, it is concluded that conservative estimation of output PDF is essential to build a safe design when insufficient experimental data is provided. Despite that the increased cost by a conservatively designed product, only few actual product testing can be used, which may lead to cost-effectiveness overall. Thus, the user can have trade-off option between the testing cost and optimized product design.

Table 4. Estimated Probability of Failure of Updated Simulation Model under Different Data Sizes
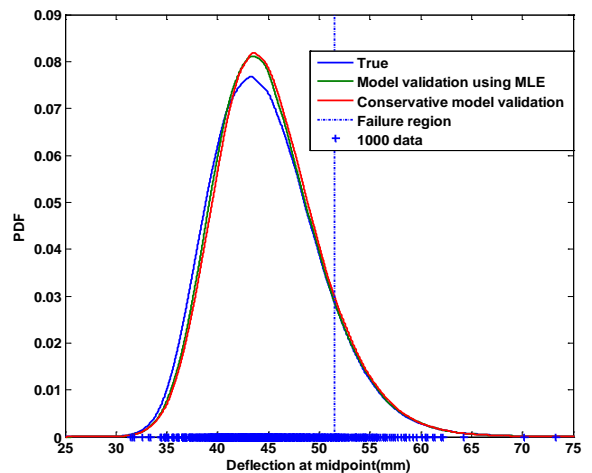
Data size	Conservative model validation	Model validation using MLE	True P_F
5	18.90%	9.18%	11.09%
10	9.44%	5.35%	11.09%
50	11.01%	7.75%	11.09%
100	12.80%	10.96%	11.09%
300	13.73%	13.24%	11.09%
500	12.13%	11.92%	11.09%
1000	11.68%	11.26%	11.09%
2000	11.64%	11.52%	11.09%
3000	11.98%	11.47%	11.09%
4000	11.95%	11.60%	11.09%

(a) 5 Experimental data

(b) 50 Experimental data

(c) 1000 Experimental data
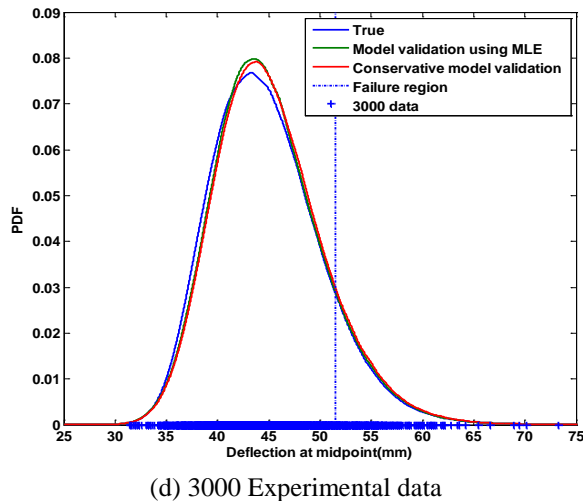
(d) 3000 Experimental data

Figure 8. Effect of Different Experimental data Sizes on Simulation Output PDF

7. DISCUSSIONS AND CONCLUSIONS

In this paper, we proposed a conservative model validation approach that assures a reliable and conservative simulation model even with a lack of experimental data. The key feature of the proposed method is that it investigates and quantifies the uncertainty induced by the limited experimental data. Under the uncertainty, predicted output PDF and probability of failure become uncertain and subjective. To tackle this issue, a conservative estimation to obtain the target output PDF and probability of failure for validation have been successfully developed using Bayesian analysis. Then a model-updating optimization has been successfully applied using the complex variable method (CVM) for sensitivity analysis to provide a conservative fit of the simulation output PDF to the conservative target output PDF. The optimization process minimizes the distance between the simulation output PDF and the conservative target output PDF by satisfying the conservative probability of failure. The principle of this proposed model validation approach is to achieve reliable and conservative simulation-based design by estimating the probability of failure of the simulation model at a user-specified conservativeness level. The result using the developed conservative model validation was compared to the result using MLE, which has been widely used in the model validation area. It was shown that the model validation using MLE fails to avoid the underestimation of the probability of failure. On the other hand, it was demonstrated that the conservative model validation proposed in this paper provides conservative and safe simulation model even with a lack of experimental data. As the simulation model is not perfect, although the characterized calibration parameter and bias may not be exactly accurate, it provides conservative fit to the true output PDF. Therefore, a product design with the simulation model, which is updated using the developed method, will be a reliable design even with a lack of experimental data. The conservative model validation

presented in this paper can be applied to a wide variety of physical systems for design purposes. Future work will extend to model validation under design changes so that it can be integrated with RBDO.

ACKNOWLEDGMENTS

REFERENCES
[1] Hills, R. G., and Leslie, I. H., 2003, "Statistical Validation of Engineering and Scientific Models: Validation Experiments to Application," SAND2003-0706, Sandia National Laboratories, Albuquerque, NM.

[2] Oberkampf, W. L., and Barone, M. F., 2006, "Measures of Agreement between Computation and Experiment: Validation Metrics," *Journal of Computational Physics*, **217**(1), pp. 5-36.

[3] Chen, W., Xiong, Y., Tsui, K. L., and Wang, S., 2006, "Some Metrics and a Bayesian Procedure for Validating Predictive Models in Engineering Design," *Proceedings of ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Philadelphia, American Society of Mechanical Engineers, pp. 1205-1214

[4] Ferson, S., Oberkampf, W. L., and Ginzburg, L., 2008, "Model Validation and Predictive Capability for the Thermal Challenge Problem," *Computer Methods in Applied Mechanics and Engineering*, **197**(29), pp. 2408-2430.

[5] Kennedy, M. C., and O'Hagan, A., 2001, "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63**(3), pp. 425-464.

[6] Higdon D., Nakhleh C., Gattiker J., and Williams B., 2008, "A Bayesian Calibration Approach to the Thermal Problem," *Computer Methods in Applied Mechanics and Engineering*, **197**(29-32), pp. 2431-2441.

[7] Arendt, P., Apley, D., and Chen, W., 2012, "Quantification of Model Uncertainty: Calibration, Model Discrepancy, and Identifiability," *Journal of Mechanical Design*, **134**(10).

[8] Arendt, P., Apley, D., Chen, W., Lamb, D., and Gorsich, D., 2012, "Improving Identifiability in Model Calibration Using Multiple Responses," *Journal of Mechanical Design*, **134**(10).

[9] Loeppky, J., Bingham, D., and Welch, W., 2006, "Computer Model Calibration or Tuning in Practice," Technical Report University of British Columbia, Vancouver, BC, CA.

[10] Xiong Y., Chen W., Tsui K. L., and Apley D., 2009, "A Better Understanding of Model Updating Strategies in Validating Engineering Models," *Computer Methods in*

Applied Mechanics and Engineering, **198**(15-16), pp. 1327-1337.

[11] Youn, B. D., Jung, B. C., Xi, Z., Kim, S. B., and Lee, W. R., 2011, "A Hierarchical Framework for Statistical Model Calibration in Engineering Product Development," *Computer Methods in Applied Mechanics and Engineering*, **200**(13-16), pp. 1421–1431.

[12] Acar, E., Kale, A., Haftka, R. T., and Stroud, W. J., 2006, "Structural Safety Measures for Airplanes," *Journal of Aircraft*, **43**(1), pp. 30-38.

[13] Picheny, V., Kim, N. H., and Haftka, R. T., 2010, "Application of Bootstrap Method in Conservative Estimation of Reliability with Limited Samples," *Structural and Multidisciplinary Optimization*, **41**(2), pp. 205-217.

[14] Youn, B. D., and Wang, P., 2008, "Bayesian Reliability-based Design Optimization using Eigenvector Dimension Reduction (EDR) Method," *Structural and Multidisciplinary Optimization*, **36**(2), pp. 107-123.

[15] Noh, Y., Choi, K.K., Lee, I., Gorsich, D., and Lamb, D., 2011, "Reliability-Based Design Optimization with Confidence Level under Input Model Uncertainty Due to Limited Test Data," *Structural and Multidisciplinary Optimization*, **43**, pp. 443–458.

[16] Noh, Y., Choi, K.K., Lee, I., and Gorsich, D., 2011, "Reliability-based Design Optimization with Confidence Level for Non-Gaussian Distributions Using Bootstrap Method," *ASME Journal of Mechanical Design*, **133**, 091001.

[17] Cho, H., 2014, "Efficient Variable Screening Method and Confidence-based Method for Reliability-based Design Optimization," Ph.D. Thesis, University of Iowa.

[18] Pearson, K., 1916, "Mathematical Contributions to the Theory of Evolution, XIX: Second Supplement to a Memoir on Skew Variation," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, **216**, pp. 429-457

[19] Martins, J. R., Sturdza, P., and Alonso, J. J., 2003, "The Complex-Step Derivative Approximation," *ACM Transactions on Mathematical Software (TOMS)*, **29**(3), pp. 245-262.

[20] Bhatti, M. A., 2006, *Advanced Topics in Finite Element Analysis of Structures: with Mathematica and MATLAB computations*, John Wiley & Sons, Inc.

[21] Zhao, L., Choi, K. K., and Lee, I., 2011, "Metamodeling Method using Dynamic Kriging for Design Optimization," *AIAA journal*, **49**(9), pp. 2034-2046.

[22] Song, H., Choi, K. K., and Lamb, D., 2013, "A study on improving the accuracy of kriging models by using correlation model/mean structure selection and penalized log-likelihood function," *10th World Congress on Structural and Multidisciplinary Optimization*, Florida, Orlando

[23] Volpi, S., Diez, M., Gaul, N. J, Song, H., Iemma, U., Choi, K. K., Campana, E. F., and Stern, F., 2014, "Development and Validation of a Dynamic Metamodel based on Stochastic Radial Basis Functions and Uncertainty Quantification," *Structural and Multidisciplinary Optimization*, pp.1 -22.

[24] Sen, O., Davis, S., Jacobs, G., and Udaykumar, H. S., 2014, "Evaluation of Convergence Behavior of Metamodeling Techniques for Bridging Scales in Multi-Scale Multimaterial Simulation." *submitted to Journal of Computational Physics*

[25] Chakrabarty, J., 2006, *Theory of Plasticity*, Elsevier Butterworth-Heinemann, Burlington, MA.